Water Resources Variability in Africa during the XXth Century (Proceedings of the Abidjan'98 Conference held at Abidjan, Côte d'Ivoire, November 1998). IAHS Publ. no. 252, 1998.

La procédure de segmentation, dix ans après

P. HUBERT
UMR Sisyphe, CIG, Ecole des Mines de Paris, 35 Rue Saint Honoré, F-77305 Fontainebleau, France
e-mail: hubert@cig.ensmp.fr

E. SERVAT, J.-E. PATUREL, B. KOUAME
Antenne Hydrologique, ORSTOM, 06 BP 1203 Cedex 1, Abidjan 06, Côte d'Ivoire

H. BENDJOUDI, J.-P. CARBONNEL
UMR Sisyphe, Laboratoire de Géologie Appliquée, Université Pierre et Marie Curie, 4 Place Jussieu, F-75252 Paris Cedex 05, France

H. LUBES-NIEL
ORSTOM, BP 5045, F-34032 Montpellier Cedex, France

Résumé La brutalité des modifications pluviométriques enregistrées en Afrique au début des années soixante-dix avait alors amené deux d'entre nous (P. H. et J.-P. C.) à nous intéresser, pour les séries chronologiques de pluies annuelles, à des modèles faisant appel à des discontinuités et en particulier à des sauts de la tendance centrale. Nous avions dans ce but défini une procédure de segmentation, permettant d'identifier une ou plusieurs discontinuités de ce type dans une série chronologique, dont l'application à une cinquantaine de séries pluviométriques Sahéliennes fut publiée en 1989. Cette étude avait permis de distinguer au cours du XXe siècle une succession de 5 phases alternativement relativement sèches et relativement humides. Presque dix ans plus tard, il nous a semblé intéressant de reprendre cette étude sur les séries des mêmes stations lorsqu'il s'est révélé possible de rassembler de nouvelles données. Une telle étude a un double intérêt: d'abord faire une nouvelle analyse de l'évolution pluviométrique au Sahel au cours du XXe siècle, mais aussi évaluer la robustesse de la procédure de segmentation. Cette nouvelle étude a pu être menée sur une vingtaine de séries. Elle montre que la phase sèche qui a débuté vers 1970 se poursuit. En ce qui concerne la procédure de segmentation elle se révèle particulièrement robuste puisque l'allongement des séries soumises à l'analyse n'a en général pas modifié l'identification des discontinuités initialement mises en évidence.

Les hydrologues, et derrière eux les aménageurs, ont longtemps considéré les séries hydropluviométriques comme stationnaires. C'est dans cet esprit qu'est née la notion de **norme** pluviométrique et/ou hydrologique (Morel, 1986) calculée sur une période de référence, fréquemment utilisée dans les calculs destinés à orienter les pratiques agricoles (Jackson, 1970) ou à concevoir l'aménagement des bassins versants. La succession dans la région soudano-sahélienne d'années sèches observées au début des années soixante-dix (Sircoulon, 1976; Dhonneur, 1981; Nicholson, 1983; Olivry, 1983) devait amener, parfois douloureusement, les hydroclimatologues à revenir sur ce credo et à imaginer pour les séries annuelles de pluies et de débits des modèles plus complexes. Si la persistance (Brunet-Moret & Roche, 1975) est une

caractéristique des séries hydrométéorologiques annuelles de l'Afrique de l'ouest, qui se traduit souvent par le rejet de l'hypothèse d'indépendance statistique et par des coefficients de Hurst élevés (Hurst, 1950; Hubert & Carbonnel, 1987), il était devenu impossible au cours des années quatre-vingts d'expliquer une période sèche dans le cadre d'un modèle stationnaire, fût il autorégressif (Snijders, 1986; Musy & Meylan, 1987). La vieille théorie des cycles a été invoquée (Gac & Faure, 1981) mais a vite fait long feu. Il faut noter qu'en l'absence de très longues séries qui permettraient de faire sérieusement de l'analyse spectrale, les chercheurs utilisent généralement la douteuse méthode des moyennes mobiles dont les artefacts (effet Slutsky) qui introduisent des périodicités artificielles ont été maintes fois dénoncés (Bernier, 1965). Compte tenu de la brutalité de certaines évolutions observées en Afrique de l'ouest, de nombreux auteurs se sont orientés vers la recherche d'une ou plusieurs singularités au sein des séries. Ces singularités peuvent en particulier consister en un changement de moyenne provoqué par une modification d'origine naturelle et/ou anthropique du processus physique de génération. A cet effet de nombreux tests et/ou procédures plus ou moins classiques, plus ou moins Bayésiens aussi, ont été utilisés et/ou élaborés pour mettre en évidence un changement de moyenne au sein d'une série chronologique. Sans prétendre à l'exhaustivité nous citerons les méthodes évoquées ou définies par Blanchet *et al.* (1971), Lee & Heghinian (1977), Pettitt (1979) et Buishand (1982). Les outils correspondants ne permettaient cependant que la recherche d'une singularité unique. Certains auteurs (Klemes, 1974; Potter, 1976) avaient évoqué, sur un plan purement théorique, des modèles faisant appel à des singularités multiples, susceptibles d'interprétations physiques simples et rendant compte du comportement du coefficient de Hurst, et c'est dans cet esprit que nous avons conçu et implémenté une procédure de segmentation de séries chronologiques.

Nous ne décrirons pas ici par le menu cette procédure qui a déjà fait par ailleurs l'objet de publications (Hubert *et al.*, 1989; Hubert, 1997). Disons que cette procédure recherche la meilleure segmentation (découpage) de la série soumise à l'analyse en sous séries contiguës, et ceci pour des ordres (nombre de segments ou sous-séries) croissants. Pour chaque ordre de segmentation, la meilleure segmentation est celle qui minimise une distance égale à la somme des écarts quadratiques entre chacune des valeurs de la série et sa moyenne locale (moyenne de la sous-série à laquelle appartient une valeur). On contrôle la pertinence d'un segmentation en vérifiant que la différence entre toutes les moyennes locales contiguës prises deux à deux est significative, ce qui est réalisé au moyen du test de Scheffe (1959), utilisant la notion de contraste dont on trouvera un exposé simplifié dans Dagnélie (1970), test pour lequel un niveau de signification devra être défini. Ce contrôle permet d'ailleurs d'arrêter le processus de segmentation qui se poursuivrait autrement jusqu'à la segmentation d'une série de n éléments en ... *n* segments, segmentation pour laquelle le critère de distance défini plus haut s'annule. Le nombre de segmentations possibles d'une série peut devenir extrêmement important et il est hors de question, à chaque ordre de segmentation, de procéder à une exploration systématique des segmentations possibles pour déterminer la segmentation optimale. Le point le plus délicat de l'algorithme de la procédure de segmentation consiste donc à éviter cette exploration systématique et l'explosion

combinatoire correspondante. Cela est réalisé en organisant les segmentations possibles selon un arbre dont l'exploration sera limitée grâce à un algorithme de type "branch and bound" (Minoux, 1983).

Depuis sa conception et ses premières applications à des séries hydropluviométriques de la zone soudano-sahélienne (Hubert *et al.*, 1989), la procédure de segmentation a connu de nombreuses applications dans des contextes variés. Chaouche (1988) l'a utilisée dans sa thèse sur la structure de la saison des pluies en Afrique Soudano-sahélienne. D'autres applications en Afrique de l'ouest et centrale ont été réalisées par Moron (1992), Lubès *et al.* (1995), Paturel *et al.* (1997a, 1997b), Servat *et al.* (1997a, 1997b) ou Laraque *et al.* (1997) et nous avons nous-même *revisité* après quelques années notre étude sur le débit des grands fleuves Africains (Hubert & Carbonnel, 1993). Kebaili-Bargaoui (1990) s'en est servi sur des séries Tunisiennes. Hors d'Afrique, la procédure de segmentation a été employée par Slivitsky & Mathier (1993) dans une étude sur l'évolution climatique de la région des grands lacs Nord-Américains et elle a été appliquée par Carbonnel & Hubert (1994) et Carbonnel *et al.* (1994) à l'étude de séries pluviométriques Bulgares et Roumaines. D'autres applications concernant des données d'Amérique du Sud ont été publiées (Ronchail, 1996). Par ailleurs la procédure de segmentation a été évaluée par Bernier (1993) qui a indiqué que les points de rupture déterminés par cette procédure étaient optimaux selon un point de vue Bayésien. Cavadias (1992, 1993) l'a citée dans son enquête menée sous l'égide de l'OMM sur les outils disponibles pour l'étude des modifications climatiques. Signalons enfin que la procédure de segmentation a été intégrée au logiciel KhronoStat développé au sein du Programme Grands Bassins Tropicaux de l'ORSTOM (Lubès *et al.*, 1994; Ecole des Mines de Paris et l'UMR GBE de l'Université de Montpellier II, 1998).

Dans notre article de 1989, nous avions étudié 33 séries pluviométriques ainsi que les chroniques du fleuve Sénégal à Bakel et du fleuve Niger à Koulikoro. La fin des séries étudiées s'établissait vers 1982. Une dizaine d'années plus tard, il nous a semblé intéressant de reprendre cette étude, avec des séries un peu plus longues. Il ne nous a été possible de compléter de façon notable que 21 séries (Tableau 1) provenant de stations du Sénégal, du Mali, du Burkina Faso et du Niger. Ces séries s'achèvent généralement vers 1993, soit un gain de 12 ans sur la longueur moyenne de ces séries. La longueur moyenne des séries constituant le corpus est donc maintenant de 66 ans (au lieu de 49 pour le corpus étudié en 1989).

On trouvera sur la Fig. 1 une synthèse des résultats de la segmentation de ces séries. Les segmentations ont été réalisées pour le niveau de signification 0.01 du test de Scheffe (limitant le risque de segmenter une série qui ne devrait pas l'être qui est ici un risque de première espèce), les segmentations conduisant à des segments jugés trop courts (inférieurs à 5 ans) ayant été rejetées. On constate un grand nombre (20) de ruptures négatives détectées à la fin des années soixante (leur moyenne pondérée conduit à retenir 1967–1968 comme date de transition). On observe de plus trois ruptures positives au début des années cinquante et trois ruptures négatives au début des années quarante. Assez curieusement on observe une rupture positive isolée pour la transition 1935–1936 (elle est observée pour la série de Bougouni au Mali qui semble très atypique car elle présente aussi la particularité d'être la seule série du corpus à ne pas présenter de rupture négative à la fin des années soixante). Les

Tableau 1 Liste des stations utilisées dans cette étude (On a indiqué leur nom, le pays auquel elles appartiennent, leur coordonnées géographiques, dans la colonne [JoH] l'étendue de la série utilisée par Hubert *et al* (1989) et dans la colonne [Etat 98] l'étendue de la série utilisée dans la présente étude).

	Station		Lat.	Long.	JdH 89	Etat 98
1	Dakar-Yoff	SN	14,44	-17,30	1947-1982	1947-1994
2	Kaolack	SN	14,08	-16,04	1921-1982	1921-1994
3	Matam	SN	15,38	-13,15	1936-1982	1927-1994
4	Podor	SN	16,39	-14,56	1930-1978	1928-1985
5	Sedhiou	SN	12,42	-15,33	1938-1979	1927-1987
6	Bamako_Aero	M	12,38	-8,01	1929-1982	1929-1992
7	Bougouni	M	11,25	-7,20	1922-1982	1921-1994
8	Kayes	M	14,26	-11,26	1943-1982	1934-1994
9	Kita	M	13,04	-9,27	1931-1982	1931-1994
10	Mopti	M	14,31	-4,06	1936-1982	1935-1994
11	San	M	13,17	-4,54	1922-1982	1922-1994
12	Segou	M	13,24	-6,09	1936-1982	1935-1994
13	Bobo-Dioulasso	BF	11,10	-4,18	1938-1976	1920-1994
14	Dedougou	BF	12,28	-3,29	1925-1980	1922-1994
15	Tenkodogo	BF	11,46	-0,23	1928-1979	1922-1990
16	Fada-N'Gourma	BF	12,04	-0,21	1933-1976	1933-1993
17	Kaya	BF	13,06	-1,05	1921-1972	1921-1990
18	Agades	NE	16,58	7,59	1945-1984	1943-1994
19	Birni N'Koni	NE	13,48	5,17	1934-1984	1934-1994
20	Niamey Aero	NE	13,29	2,10	1944-1984	1924-1994
21	Zinder	NE	13,47	8,59	1922-1984	1911-1994
	Senegal a Bakel	SN	14,54	-12,38	1903-1986	1903-1993
	Niger a Koulikoro	M	12,58	-7,22	1907-1986	1907-1994

résultats sont donc très comparables à ceux que nous avions obtenus il y a dix ans et qui ont été souvent confirmés par les travaux que nous avons cités plus haut. Ils indiquent tous que la sécheresse qui a débuté fin des années soixante—début des années soixante-dix se poursuit encore à l'heure actuelle puisqu'aucune segmentation positive n'a été enregistrée depuis.

Cette permanence est particulièrement bien illustrée par la segmentation des séries de débits moyens annuels du fleuve Sénégal à Bakel et du fleuve Niger à Koulikoro. Le comportement de ces deux séries est très similaire (le coefficient de corrélation de leur partie commune 1907-1993 est égal à 0.727). Les résultats des deux segmentations sont représentés sur la Fig. 2. Toutes les ruptures observées dans notre article de 1989 se retrouvent dans l'analyse réalisée de nos jours (à savoir 1921-1922 [+], 1936-1937 [-], 1949-1950 [+] et 1967-1968 [-] pour le fleuve Sénégal, 1923-1924 [+], 1932-1933 [-], 1950-1951 [+] et 1969-1970 [-] pour le fleuve Niger). La seule innovation apportée par 7 ou 8 années supplémentaires est la détection d'une nouvelle rupture (négative) lors de la transition 1979-1980 sur la série du fleuve Niger.

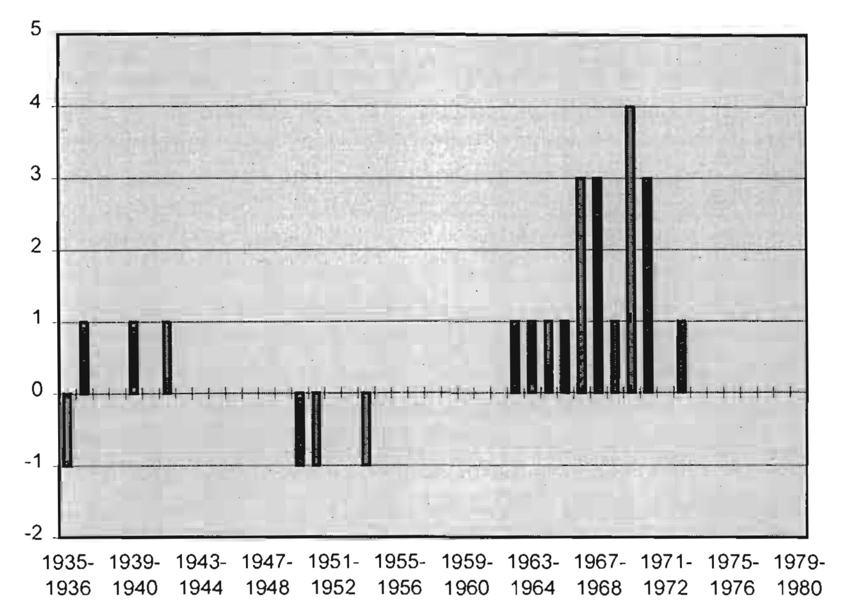


Fig. 1 Histogramme des 26 ruptures observées sur les 21 séries du corpus. Les ruptures positives ont été portées vers le *bas* et les ruptures négatives vers le *haut*.

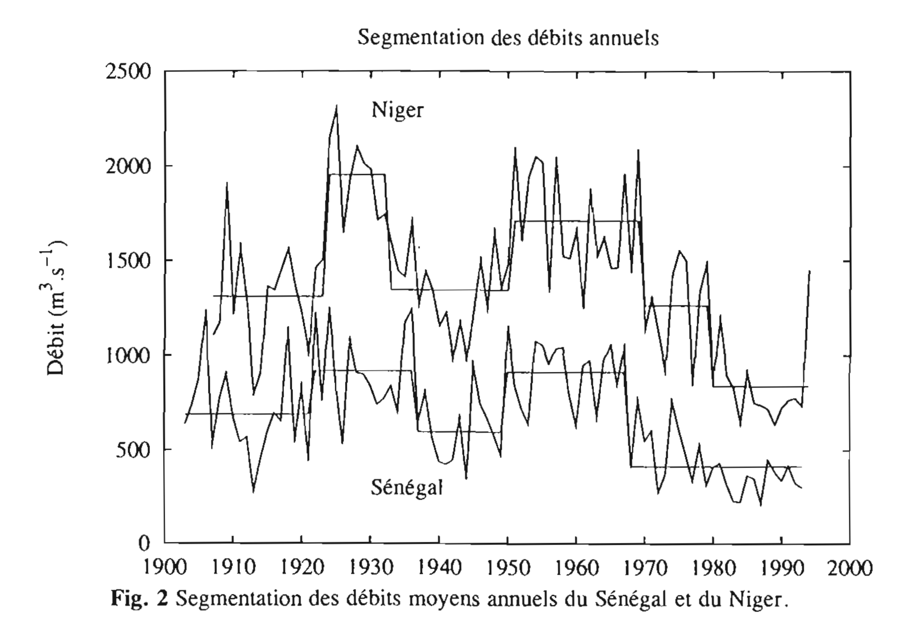


Fig. 2 Segmentation des débits moyens annuels du Sénégal et du Niger.

CONCLUSIONS

Nous ne nous apesantirons pas sur la signification climatologique de ces résultats. Ils ne font que confirmer des résultats aujourd'hui bien connus concernant la succession au cours du XXème siècle de diverses phases climatiques contrastées en Afrique de

l'ouest. Ces conclusions s'appuient sur des analyses indépendantes et concordantes de nombreuses séries de la région. Il nous semble plus important ici de souligner la robustesse et la fiabilité de la procédure de segmentation qui se révèle peu sensible aux effets de bord. Ces qualités, confirmées dans de nombreuses études menées sur des séries d'origines et de caractéristiques variées, devraient justifier de continuer à développer cette procédure. Il semble possible d'améliorer les performances de l'algorithme et donc de diminuer les temps de calcul. Cela permettrait d'envisager d'adapter et d'étendre l'algorithme au traitement de séries présentant des données manquantes. La disponibilité d'un tel outil autoriserait l'utilisation d'une matériel considérable aujourd'hui très peu valorisable pour les études d'évolution climatique.

Le programme de la procédure de segmentation (module exécutable fonctionnant sous MS-DOS), son mode d'emploi (en français, anglais et portugais), ainsi qu'un jeu de données type, sont disponibles en libre service sur Internet à l'adresse suivante:

http://www.cig.ensmp.fr/~hubert

REFERENCES

Bernier, J. (1965) Sur les probabilités d'occurrence des sécheresses et des étiages. *Bull. Cent. Rech. Essais de Chatou* **11**, 3-12.

Bernier, J. (1993) Statistical detection of changes in geophysical series. Communication to the NATO Advanced Study Institute *Engineering Risk and Reliability in a Changing Physical Environment* (Deauville, France, 24 May-4 June 1993).

Blanchet, M., Bouzid, M., Delhomme, R., Ouakam, J. P., Pietri, M. & Bernier, J. (1971) Phénomènes de rupture dans une série chronologique. *Rev. Stat. Appl.* **19**(4), 5-22.

Brunet-Moret, Y. & Roche, M. (1975) Persistance dans les suites chronologiques de précipitations annuelles. *Cah. ORSTOM, Série Hydrol.* **12**, 147-165.

Buishand, T. A. (1982) Source methods for testing the homogeneity of rainfall records. *J. Hydrol.* **58**, 11-27.

Carbonnel, J. P. & Hubert, P. (1994) Etude statistique de quelques séries pluviométriques roumaines et bulgares. Incidences sur l'évolution climatique récente, Romanian . *J. Hydrol. & Wat. Resour.* **1**, 107-113.

Carbonnel, J. P., Mihnevski, N. & Hubert, P. (1994) L'évolution récente de la pluviométrie dans le massif de Rila. *Observatoire de montagne de Moussala, Sofia* **2**, 52-58.

Cavadias, G. (1992) A survey of current approaches to modelling of hydrological time series with respect to climate variability and change. Communication to the WMO Workshop on WCP—Water(Geneva, 3-13 November 1992).

Cavadias, G. (1993) Detection and modelling of the impact of climatic change on River Flows. Communication to the NATO Advanced Study Institute *Engineering Risk and Reliability in a Changing Physical Environment* (Deauville France, 24 May-4 June 1993).

Chaouche, A. (1988) Structure de la saison des pluies en Afrique Soudano-sahélienne. Thèse de doctorat de l'Ecole des Mines de Paris, Paris, France.

Dagnélie, P. (1970) *Théorie et Méthodes Statistiques*, vol. 2. Duculot, Gembloux, France.

Dhonneur, G. (1981) Sécheresse et/ou désertification au Sahel. *La Météorologie, Paris, Série VI* **24**, 119-123.

Ecole des Mines de Paris et l'UMR GBE de l'Université de Montpellier II (1998)Logiciel associé: KhronoStat version 1.0.

Faure, M. & Gac, J. Y. (1981) Will the Sahelian drought end in 1985? *Nature* **291**(5815), 475-478.

Hubert, P. (1997) Change points in hydrometeorological time series. In: *Proc. Conf. Applications of Time Series Analysis in Astronomy and Meteorology*, 399-412. Chapman & Hall, UK.

Hubert, P. & Carbonnel, J. P. (1987) Approche statistique de l'aridification de l'Afrique de l'ouest. *J. Hydrol.* **95**, 165-183.

Hubert P. & Carbonnel, J. P. (1993) Segmentation des séries annuelles de débits de grands fleuves Africains. *Bull. du CIEH, Ouagadougou* **92**, 3-10.

Hubert, P., Carbonnel, J. P. & Chaouche, A. (1989) Segmentation des séries hydrométéorologiques. Application à des séries de précipitations et de débits de l'Afrique de l'ouest. *J. Hydrol* **10**, 349-367.

Hurst, H. E. (1950) Long-term storage capacity of reservoirs. *Proc. Am. Soc. Civ. Engrs* **76**(11).

Jackson, I. J. (1970) Annual rainfall probability and the binomial distribution. *East Afr. For. J.* **36**, 265-272.

Kebaili-Bargaoui, Z. (1990) Modélisation stochastique des sécheresses annuelles et pluriannuelles. Thèse de doctorat d'état, Faculté des Sciences de Tunis, Tunis, Tunisie.

Klemes, V. (1974) The Hurst phenomena: a puzzle ? *Wat. Resour. Res.* **10**, 675–688.

Laraque, A., Olivry, J. C., Orange, D. & Marieu, B. (1997) Variations spatio-temporelles des régimes pluviométriques et hydrologiques en Afrique centrale du début du siècle à nos jours. In: *FRIEND'97—Regional Hydrology: Concepts and Models for Sustainable Water Resource Management* (ed. by A. Gustard *et al.*) (Proc. FRIEND Conf., Postojna, Slovenia, September–October 1997), 257–266. IAHS Publ. no. 246.

Lee, A. F. S. & Heghinian, S. M. (1977) A shift of the mean level in a sequence of independant normal random variable: a Bayesian approach. *Technometrics* **19**, 503–506.

Lubès, H., Masson, J. M., Servat, E., Paturel, J. E., Kouame, B. & Boyer, J. F. (1994) *Caractérisation de fluctuations dans une série chronologique par application de tests statistiques. Etude bibliographique.* Programme ICCARE Rapport no.°3. ORSTOM, Montpellier, Côte d'Ivoire, juin 1994.

Lubès, H., Aka, A., Masson, J. M., Servat, E., Paturel, J. E. & Kouamé, B. (1995) Essai de mise en évidence d'une variation climatique par application de tests statistiques à des séries chronologiques de débit. Application aux grands fleuves de Côte d'Ivoire, à paraître dans les C.R. de Méthodes Statistiques et Bayésiennes en Hydrologie. Proc. Conférence Internationale en l'honneur de Jacques Bernier.

Minoux, M. (1983) *Programmation Mathématique. Théorie et Algorithmes*, tome 2. Dunod, Paris, France.

Morel, R. (1986) Problèmes posés par les normes pluviométriques dans la région sahélienne. Colloq. Int. Normes Hydrol. Suite aux incidences de la sécheresse, CIEH, Ouagadougou.

Moron, V. (1992) Variabilité spatio-temporelle des précipitations en Afrique sahélienne et guinéenne de 1933 à 1990. *La Météorologie, Paris, Série VII* **43–44**, 24–30.

Musy, A. & Meylan, P. (1987) Modélisation d'un processus non-stationnaire: application à la pluviométrie en zone semi-aride. In: *The Influence of Climatic Change and Climatic Variability on the Hydrologic Regime and Water Resources* (ed. par S. I. Solomon, M. Beran & W. Hogg) (Proc. Vancouver Symp., August 1987), 287–299. IAHS Publ. no. 168.

Nicholson, S. E. (1983) Sub-Saharan rainfall in the years 1976–1980: evidence of continued drought. *Mon. Weath. Rev.* **111**, 1646–1654.

Olivry, J. C. (1983) Le point en 1982 sur l'évolution de la sécheresse en Sénégambie et aux Iles du Cap Vert. Examen de quelques séries de longue durée (débits et précipitations). *Cah. ORSTOM, Série Hydrol.* **20**, 47–69.

Paturel, J. E., Kouamé, B., Travaglio, M., Lubès, H., Marieu, B., Fritsch, J. M. & Masson, J. M. (1997a) Modification des régimes d'écoulement en Afrique de l'ouest et centrale non sahélienne et conséquences sur les ressources en eau. In: *FRIEND'97-Regional Hydrology: Concepts and Models for Sustainable Water Resource Management* (ed. by A. Gustard *et al.*) (Proc. FRIEND Conf., Postojna, Slovenia, September–October 1997), 241–248. IAHS Publ. no. 246.

Paturel, J. E., Servat, E., Kouame, B., Lubès, H., Ouedraogo, M. & Masson, J. M. (1997b) Climatic variability in humid Africa along the Gulf of Guinea, Part 2: An integrated regional approach. *J. Hydrol.* **191**, 16–36.

Pettitt, A. N. (1979) A non-parametric approach to the change-point problem. *Appl. Stat.* **28**, 126–135.

Potter, K. W. (1976) Evidence of nonstationarity as a physical explanation of the Hurst phenomenon. *Wat. Resour. Res.* **12**, 1047–1052.

Ronchail, J. (1996) Variabilité pluridécennale des précipitations en Bolivie: essai de mise en relation avec les températures de surface océaniques de l'Atlantique extratropical. *Publications de l'Association Internationale de Climatologie* **9**, 504–511.

Scheffe, M. (1959) *The Analysis of Variance.* Wiley, New York, USA.

Servat, E., Paturel, J. E., Lubès, H., Kouame, B., Ouedraogo, M. & Masson, J. M. (1997a) Climatic variability in humid Africa along the Gulf of Guinea. Part 1: Detailed analysis of the phenomenon in Côte d'Ivoire. *J. Hydrol.* **191**, 1–15.

Servat, E., Paturel, J. E., Lubès, H., Kouame, B. & Masson, J. M. (1997b) Variabilité des régimes pluviométriques en Afrique de l'ouest et centrale non sahélienne. *C.R. Acad. Sci., Paris* **324**, série IIa, 835–838.

Sircoulon, J. (1976) Les données hydrométéorologiques de la sécheresse récente en Afrique Intertropicale. Comparaison avec les sécheresses 1913 et 1940. *Cah. ORSTOM, Série Hydrol.* **13**, 75–174.

Slivitsky, M. & Mathier, L. (1993) Climatic changes during the 20th century on the Laurentian Great Lakes and their impacts on hydrological regime. Communication to the NATO Advanced Study Institute *Engineering Risk and Reliability in a Changing Physical Environment* (Deauville, France, 24 May–4 June 1993).

Snijders, T. A. (1986) Interstation correlations and non stationarity of Burkina Faso rainfall. *J. Clim. Appl. Met.* **25**, 524–531.